

**ĐẠI HỌC QUỐC GIA HÀ NỘI
TRƯỜNG ĐẠI HỌC CÔNG NGHỆ**



Nguyễn Tuấn Anh

**PHÁT TRIỂN PHẦN MỀM NHẬN DẠNG HOA
TRÊN NỀN TẢNG THIẾT BỊ DI ĐỘNG**

**KHÓA LUẬN TỐT NGHIỆP ĐẠI HỌC HỆ CHÍNH QUY
Ngành: Công nghệ thông tin**

HÀ NỘI - 2019

**ĐẠI HỌC QUỐC GIA HÀ NỘI
TRƯỜNG ĐẠI HỌC CÔNG NGHỆ**

Nguyễn Tuấn Anh

**PHÁT TRIỂN PHẦN MỀM NHẬN DẠNG HOA
TRÊN NỀN TẢNG THIẾT BỊ DI ĐỘNG**

KHÓA LUẬN TỐT NGHIỆP ĐẠI HỌC HỆ CHÍNH QUY
Ngành: Công nghệ thông tin

Cán bộ hướng dẫn: TS. Nguyễn Thị Ngọc Diệp

Cán bộ đồng hướng dẫn: PGS.TS. Nguyễn Việt Hà

HÀ NỘI - 2019

**VIETNAM NATIONAL UNIVERSITY, HA NOI
UNIVERSITY OF ENGINEERING AND TECHNOLOGY**

Nguyen Tuan Anh

**DEVELOPING A FLOWER RECOGNITION
SOFTWARE ON MOBILE DEVICE PLATFORM**

**BACHELOR'S THESIS
Major: Information Technology**

Supervisor: Dr. Nguyen Thi Ngoc Diep

Co-Supervisor: Assoc. Prof. Nguyen Viet Ha

HANOI - 2019

LỜI CẢM ƠN

Lời đầu tiên cho tôi xin được gửi lời cảm ơn chân thành và sâu sắc nhất tới TS. Nguyễn Thị Ngọc Diệp, PGS. TS. Nguyễn Việt Hà, ThS. Nguyễn Ngọc Khuê những người đã hướng dẫn và chỉ bảo tận tình nhất cho tôi trong suốt quá trình hoàn thành khóa luận tốt nghiệp này.

Tôi xin được gửi lời cảm ơn tới toàn bộ các thầy giáo, cô giáo của trường Đại học Công Nghệ - Đại học Quốc Gia Hà Nội nhưng người đã tạo điều kiện tốt nhất để tôi có thể học tập, nghiên cứu và hơn cả là đã truyền thụ cho tôi những hành trang kiến thức đầy đủ nhất.

Tôi cũng xin gửi lời cảm ơn chân thành nhất tới các thành viên trong nhóm nghiên cứu SkyLab, tới những người bạn người anh, chị đã giúp đỡ tôi hoàn thiện cả về kiến thức chuyên môn và kỹ năng học tập nghiên cứu.

Cuối cùng và không thể thiếu đó là lời cảm ơn tới bố mẹ và chị tôi và đặc biệt là bạn Dung Phùng những người đã luôn bên cạnh tôi giúp đỡ và động viên cổ vũ tinh thần tôi trong những lúc khó khăn nhất.

Tôi xin chân thành cảm ơn!

Hà Nội, ngày 24 tháng 04 năm 2019
Sinh viên

Nguyễn Tuấn Anh

TÓM TẮT

Khóa luận này trình bày một ứng dụng nhận dạng tên loài hoa dựa vào ảnh đầu vào trên nền tảng di động. Ứng dụng này được xây dựng với ba mô hình học máy chính: (1) một mô hình phân loại nhị phân để phát hiện có đổi t hoa trong ảnh hay không, (2) một mô hình nhận dạng tên của loài hoa trong ảnh, và (3) một thuật toán tìm kiếm ảnh tương tự khi hiển thị các ảnh kết quả.

Ứng dụng được phát triển sử dụng phương pháp sử dụng lại một mô hình học sâu đã được huấn luyện trên bộ dữ liệu lớn khác để trích xuất dữ liệu của ảnh. Thực nghiệm so sánh phương pháp tiếp cận này với các cách trích xuất thuộc tính truyền thống chỉ ra tính khả thi và tiết kiệm của việc sử dụng lại mô hình giữa các bài toán nghiên cứu trong lĩnh vực xử lý ảnh.

Những đóng góp chính của khóa luận là: (1) Việt hoá tên cũng như đặc điểm về sinh trưởng, cách trồng của 102 loài hoa trong bộ dữ liệu Oxford-102, (2) phát triển ứng dụng trên nền tảng di động, và (3) phát triển mô hình phân loại ảnh có hoa hay không với độ chính xác 98.6% và mô hình nhận dạng tên loài hoa với độ chính xác là 78.56%.

Từ khóa: nhận diện hoa, phát hiện hoa, ứng dụng di động

ABSTRACT

This thesis presents an recognition flowers' names application based on the input images on the mobile platform. This application is built with three main machine learning models: (1) a binary classification model to detect whether there is any flower in the image, (2) a model that recognizes the name of the flower in the image, and (3) an algorithm which searches for results of similar images.

The application is developed using a method of reusing a Deep learning models trained by another large data set to extract feature of image. Experimental comparison of this approach with traditional feature extraction method points to the feasibility and savings of reusing model of research problems in the field of image processing.

The main contributions of this thesis are: (1) Translating names, the growth characteristics as well as planting methods of 102 flower species in Oxford-102 data set, (2) Developing an application on mobile platform, and (3) Developing a flowers detection models with 98.6% accuracy and an recognition model of flowers with 78.56% accuracy.

Keywords: *flower detection, flower recognition, mobile application*

LỜI CAM ĐOAN

Tôi xin cam đoan toàn bộ khóa luận về ứng dụng phát triển phần mềm nhận dạng hoa trên thiết bị di động bao gồm mô hình nhận diện, phần mềm di động và phần mềm hệ thống là do tôi thực hiện dưới sự hướng dẫn của TS. Nguyễn Thị Ngọc Diệp và PGS. TS. Nguyễn Việt Hà. Tất cả các công trình nghiên cứu, bài báo, khóa luận, tài liệu của các tác giả khác được tôi sử dụng trong khóa luận này đều được trích dẫn tường minh về tác giả và đều có trong danh sách tài liệu tham khảo.

Hà Nội, ngày 24 tháng 04 năm 2019
Sinh viên

Nguyễn Tuấn Anh

Mục lục

Danh sách bảng	vii
Danh sách ký hiệu, chữ viết tắt	viii
Danh sách hình vẽ	ix
Chương 1 Mở đầu	1
1.1 Bối cảnh nghiên cứu	1
1.2 Mô tả bài toán nhận dạng hoa	2
1.3 Mục tiêu và phương pháp giải quyết vấn đề	4
1.3.1 Mục tiêu của khóa luận	4
1.3.2 Phương pháp giải quyết bài toán nhận dạng hoa	4
1.3.3 Những điểm mạnh của phương pháp này	5
1.4 Bố cục các phần trong khóa luận	5
Chương 2 Nghiên cứu liên quan đến bài toán nhận dạng hoa	7
2.1 Các bộ dữ liệu hoa	7
2.2 Các nghiên cứu phân loại hoa sử dụng bộ hoa của Oxford-102	8
2.2.1 Các nghiên cứu về tiền xử lý ảnh trước khi trích xuất thông tin	8
2.2.2 Các nghiên cứu về nhận dạng dựa trên vector đặc trưng	10
2.2.3 Nghiên cứu về trích xuất thông tin từ ảnh gốc	12

2.2.4	Nghiên cứu về nhận dạng ảnh dựa trên tìm kiếm	15
2.3	Các ứng dụng di động cho nhận dạng hoa	17
Chương 3	Phát triển phần mềm nhận dạng hoa trên thiết bị di động	20
3.1	Mô tả bài toán	21
3.2	Mô tả mô hình hệ thống	22
3.2.1	Mô hình chung	22
3.2.2	Mô hình phát hiện hoa	26
3.2.3	Mô hình phân loại hoa	28
Chương 4	Kết quả thực nghiệm	30
4.1	Xây dựng bộ dữ liệu.	30
4.1.1	Bộ dữ liệu dùng để phát hiện đối tượng hoa trong ảnh.	30
Chương 5	Kết luận	32

Danh sách bảng

2.1	Bảng kết quả của nghiên cứu [2]	10
2.2	Bảng kết quả của nghiên cứu [4]	13

Danh sách ký hiệu, chữ viết tắt

MRF	Markov Random Field
HSV	Hue, saturation, value
SIFT	scale-invariant feature transform
HOG	Histogram of Gradients
SVM	Support-vector machine
SVC	Support Vector Classification
LinearSVC	Linear Support Vector Classification
KNN	K-Nearest Neighbour
RGB	Red, green, blue
CNN	Convolutional Neural Network
CNNaug	Convolutional Neural Network + image augmentation
ONE	Online Nearest-neighbor Estimation
PCA	Principal Component Analysis

Danh sách hình vẽ

1.1	Ví dụ về tìm kiếm hình ảnh của Google Images	2
1.2	Hình ảnh ví dụ về các loài hoa có cấu trúc giống nhau	3
1.3	Hình ảnh ví dụ về các loài hoa thay đổi hình dạng theo thời gian . . .	3
2.1	Hình ảnh trong nghiên cứu [1] về phương pháp tách phông nền	9
2.2	Hình ảnh trong nghiên cứu [1] về các loài hoa có hình dáng tương tự nhau	11
2.3	Hình ảnh trong nghiên cứu [1] về các loài hoa có hình dáng tương tự nhau	12
2.4	Bảng mô tả cài đặt cho mạng VGG19 trong nghiên cứu [16]	14
2.5	Hình ảnh mô tả kiến trúc của VGG19 lấy từ trang www.researchgate.net	15
2.6	Hình ảnh tóm tắt về phương pháp ước lượng láng giềng gần nhất trong nghiên cứu [3] của tác giả	16
2.7	Hình ảnh đại diện của các ứng dụng	17
2.8	Hình ảnh kết quả toàn tiếng Anh khi tra cứu ảnh hoa hướng dương bằng các ứng dụng	18
2.9	Hình ảnh về ứng dụng PlantFinder cần phải khoanh vùng phần có hoa bằng tay	19
2.10	Hình ảnh về ứng dụng chưa có chức năng phát hiện hoa trong ảnh . .	19

3.1	Ví dụ về đầu vào và đầu ra của bài toán	21
3.2	mô hình chung của hệ thống	22
3.3	ví dụ về bước tách phông nền	23
3.4	Mô hình phát hiện hoa	26
3.5	Mô hình phân loại hoa	28
4.1	Hình ảnh lấy trong bộ ảnh không có hoa	31

Chương 1

Mở đầu

Bối cảnh nghiên cứu

Cùng với sự phát triển của công nghệ nói chung cũng như sự phát triển của các thiết bị di động nói riêng ta có thể nhận thấy một điều rằng ngày nay điện thoại di động là một dụng cụ luôn đi kèm mỗi người mỗi khi đi ra khỏi nhà, đi làm hay đi chơi. Một ví dụ như sau trong một lần đi dạo bạn bỗng nhìn một bông hoa rất đẹp tròng ở ven đường và tự hỏi tên của loài hoa đó là gì. Bạn thấy rằng bông hoa đó rất quen thuộc nhưng không thể nhớ ra được. Bạn nghĩ ngay tới việc tìm kiếm trên Google và có thử lấy điện thoại di động của mình ra chụp ảnh bông hoa đó và tìm trên Google nhưng không thật may các kết quả tìm kiếm chỉ là các hình ảnh tương tự với ảnh mà bạn chụp và tên các loài hoa thì toàn là tiếng Anh và tiếng Latin, không những thế đôi khi các kết quả trả về từ Google lại không hề liên quan tới tên loài hoa đó. Qua ví dụ trên ta có thể nhận thấy nhu cầu về tìm kiếm tên của các loài hoa từ ảnh chụp thực tế từ các thiết bị di động là hoàn toàn có cơ sở. Do vậy cần một ứng dụng có chức năng nhận dạng tên loài hoa từ các ảnh chụp trên thiết bị di động.

[Floribunda \(rose\) - Wikipedia](https://en.wikipedia.org/wiki/Floribunda_(rose))

[https://en.wikipedia.org/wiki/Floribunda_\(rose\)](https://en.wikipedia.org/wiki/Floribunda_(rose)) ▾

Floribunda (Latin for "many-flowering") is a modern group of garden roses, developed by crossing hybrid teas with polyantha roses, the latter being ...

[15 Floribunda Roses for the Flower Garden -](https://www.thespruce.com/15-floribunda-roses-for-the-flower-garden-110100)

[https://www.thespruce.com > Outdoors & Gardening > Gard](https://www.thespruce.com/15-floribunda-roses-for-the-flower-garden-110100)

Feb 28, 2019 - Gene Boerner and Betty Prior Floribunda Rose palette of the hybrid tea rose, and you get floribunda roses.

About 25,270,000,000 results (1.98 seconds)



Image size:
225 × 225

Find other sizes of this image:
[All sizes](#) - [Medium](#)

Possible related search: [nepal republic day 2017](#)

Visually similar images



[Nepal Republic Day 2017 - Google](https://www.google.com/search?q=nepal+republic+day+2017)

[www.google.com/doodles/nepal-republic-day-2017](https://www.google.com/search?q=nepal+republic+day+2017) ▾

May 29, 2017 - The Republic of Nepal is a nation of immense diversity and beauty. It is the only country in the world whose flag is not rectangular!

[Nepal Republic Day 2017 - YouTube](https://www.youtube.com/watch?v=i9eyPcvvN1k)

<https://www.youtube.com/watch?v=i9eyPcvvN1k>

May 28, 2017 - Nepal Republic Day 2017. Doodle Videos. Loading... Please try again later. Published on May 28, 2017 ...

Hình 1.1: Ví dụ về tìm kiếm hình ảnh của Google Images

Mô tả bài toán nhận dạng hoa

Mô tả bài toán: bài toán nhận dạng hoa thông qua ảnh chụp sẽ có đầu vào là một hình ảnh được ghi lại từ các thiết bị ghi hình và đầu ra có thể có hai trường hợp

- Trường hợp 1: Trong ảnh không có hoa thì kết quả trả về là không tồn tại hoa.
- Trường hợp 2: Trong ảnh có hoa thì kết quả trả về là một danh sách 5 loài hoa có kết quả nhận dạng tốt nhất.

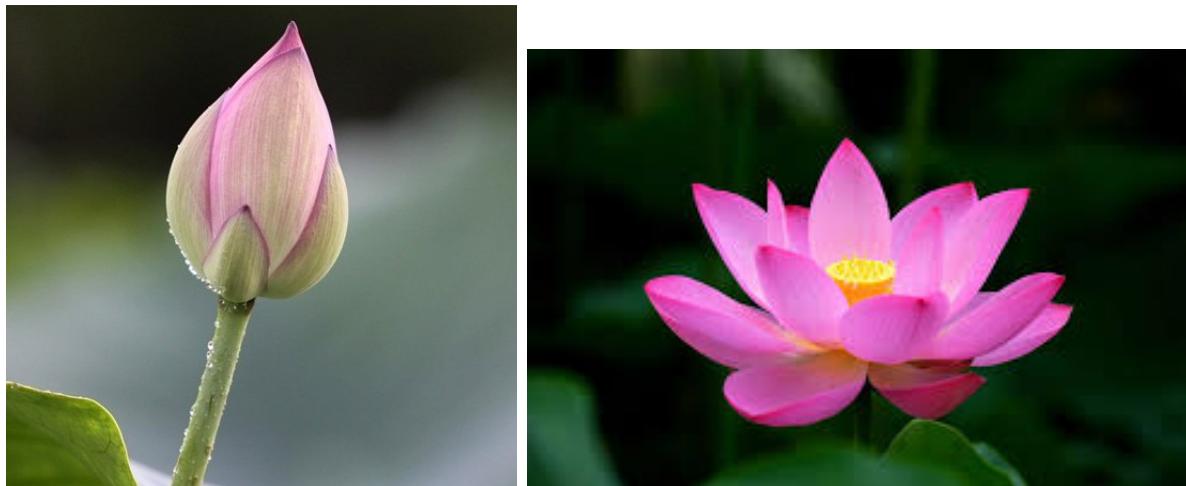
Những khó khăn khi nhận diện hoa

- Số lượng các loài hoa khác nhau mọc trong tự nhiên vô cùng đa dạng và phong phú. Các loài hoa có vô số màu sắc, kiểu dáng hình thù khác nhau gây khó khăn khi phải phân loại với một số lượng loài lớn.

- Các loài hoa đôi khi có kiểu dáng tương đối giống nhau.
- Các loài hoa trong thời gian sinh trưởng có rất nhiều hình thái khác nhau gây ra nhầm lẫn một ví dụ điển hình có thể thấy đó là hoa sen khi chưa nở và lúc đã nở có hình dạng rất khác nhau.



Hình 1.2: Hình ảnh ví dụ về các loài hoa có cấu trúc giống nhau
(Ở bên trái là ảnh hoa hướng dương còn ở bên phải là ảnh hoa cúc)



Hình 1.3: Hình ảnh ví dụ về các loài hoa thay đổi hình dạng theo thời gian
(Ở bên trái là ảnh hoa sen khi chưa nở còn ở bên phải là ảnh hoa sen khi đã nở)

Mục tiêu và phương pháp giải quyết vấn đề

Mục tiêu của khóa luận

Mục đích chính của khóa luận này muốn nhằm tới đó là phát triển một ứng dụng nhận dạng tên các loài hoa dành cho người Việt trên nền tảng thiết bị di động. Qua đó cũng đề xuất phương pháp sử dụng mạng học sâu đã được huấn luyện VGG19 [16] trong các mô hình trích xuất thuộc tính từ ảnh, đồng thời tạo tiền đề phát triển cho những nghiên cứu sau này về đề tài nhận diện ảnh.

Ứng dụng sau khi được phát triển sẽ không đơn thuần chỉ là một công cụ tra cứu mà còn có ý nghĩa trở thành một công cụ học tập, nghiên cứu về thiên nhiên và đặc biệt là các loài hoa giúp các em học sinh, mọi người ở mọi lứa tuổi có thêm một công cụ hữu ích trong việc học tập và nghiên cứu về các loài hoa.

Phương pháp giải quyết bài toán nhận dạng hoa

Để giải quyết được bài toán nhận dạng tên loài hoa sử dụng ảnh ta cần phải qua năm bước chính đó là:

- **Bước 1:** Tiền xử lý dữ liệu hình ảnh.
- **Bước 2:** Trích xuất các thuộc tính từ ảnh và chuyển thành các vector thuộc tính.
- **Bước 3:** Sử dụng các thuật toán phân loại để phát hiện có hoa trong ảnh hay không.
- **Bước 4:** Nếu phát hiện thấy hoa trong ảnh thì sử dụng vector thuộc tính đó để nhận dạng tên loài hoa
- **Bước 5:** Tính toán độ tương tự của các ảnh trong từng loài hoa để đưa ra các ảnh phù hợp nhất cho kết quả.

Qua sự tham khảo các phương pháp trích xuất truyền thống như đã được công bố trong các nghiên cứu [1] [2] và các nghiên cứu gần đây sử dụng mạng tích chập VGG19 [16]. Phương pháp để giải quyết bài toán sẽ diễn ra với các bước sau:

- Tiền xử lý hình ảnh bằng phương pháp tách phông nền.
- Trích xuất thuộc tính của ảnh sử dụng mạng VGG19 [16]
- Xây dựng một mô hình phân loại nhị phân dùng để phát hiện có đối tượng hoa trong ảnh hay không.
- Xây dựng một mô hình nhận dạng tên của loài hoa trong ảnh.
- Và cuối cùng là một thuật toán tìm kiếm ảnh tương tự khi hiển thị các ảnh kết quả.

Những điểm mạnh của phương pháp này

Qua thực nghiệm và so sánh phương pháp trích xuất thuộc tính sử dụng mô hình học sâu VGG19 [16] với các cách trích xuất thuộc tính truyền thống có kết quả khá khả quan và có tính khả thi cao.

Hơn nữa việc sử dụng lại mô hình trích xuất thuộc tính giữa các bài toán nhận dạng sẽ tiết kiệm được đáng kể công sức cũng như cải thiện được độ chính xác phân loại.

Bố cục các phần trong khóa luận

Các phần trong khóa luận được cấu trúc như sau.

- **Chương 1:** Giới thiệu về bài toán nhận dạng hoa, mục đích và các phương pháp sử dụng để giải quyết bài toán phát triển ứng dụng nhận dạng hoa trên thiết bị di động.

- **Chương 2:** Trình bày các nghiên cứu liên quan đến bài toán nhận dạng hoa. Các phương pháp trích xuất thuộc tính truyền thống và các phương pháp mới sử dụng mạng học sâu.
- **Chương 3:** Trình bày và giải thích về mô hình tổng quát của ứng dụng cũng như cách huấn luyện các mô hình phân loại được sử dụng trong hệ thống.
- **Chương 4:** Mô tả về cách thu thập dữ liệu, cách cài đặt và sử dụng các thuật toán để giải quyết vấn đề đồng thời giải thích về mục đích cũng như đưa ra các số liệu thống kê về các kết quả đã đạt được trong từng thí nghiệm.
- **Chương 5:** Đưa ra các kết luận về những vấn đề đã giải quyết và những vấn đề chưa giải quyết được đồng thời giới thiệu hướng nghiên cứu để tiếp tục phát triển trong tương lai.

Chương 2

Nghiên cứu liên quan đến bài toán nhận dạng hoa

Các bộ dữ liệu hoa

Có khá nhiều các bộ dữ liệu về hoa nổi tiếng và đa dạng ta có thể kể đến như

- Bộ dữ liệu về hoa của ImageNet.
- Oxford-17 [1] bao gồm 1360 ảnh thuộc 17 loài hoa.
- Oxford-102 [1] bao gồm 8189 ảnh thuộc 102 loài hoa.
- Kaggle flower dataset: bao gồm 4242 ảnh của 5 loài hoa.
- Bộ dữ liệu về thực vật trên trang plant database <https://garden.org/plants/>.

Trong khóa luận này sử dụng bộ dữ liệu hoa Oxford-102 [1], vì các lý do sau.

- Số lượng loài nhiều, số lượng ảnh mỗi loài cũng nhiều.
- Các ảnh chụp đều là ảnh chụp cận cảnh bông hoa và đa dạng về góc chụp, ánh sáng.

- Các ảnh mỗi loài đều có số lượng loại phong phú, ví dụ ở đây hoa hồng có cả đỏ, hồng, trắng.. thay vì chỉ một loại hoa hồng đỏ.
- Có nhiều nghiên cứu sử dụng bộ dữ liệu Oxford-102 [1], cụ thể trong khóa luận có đề cập tới 4 nghiên cứu [1] [2] [3] [4].

Bộ dữ liệu đã được chia thành bộ huấn luyện bao gồm 10 ảnh mỗi loài hoa, bộ xác nhận bao gồm 10 ảnh mỗi loại và các ảnh còn lại gồm 6129 ảnh làm bộ kiểm tra.

Các nghiên cứu phân loại hoa sử dụng bộ hoa của Oxford-102

Các nghiên cứu về phân loại hoa [1] [2] [3] [4] có sử dụng bộ dữ liệu hoa Oxford-102 [1] trong đó có thể nhóm lại thành các nhóm theo các phương pháp chính mà các nghiên cứu có sử dụng.

- Tiền xử lý ảnh trước khi trích xuất thông tin.
- Nhận dạng dựa trên vector đặc trưng.
- Trích xuất thông tin từ ảnh gốc.
- Nhận dạng dựa trên tìm kiếm (retrieval-based).

Các nghiên cứu về tiền xử lý ảnh trước khi trích xuất thông tin

Dầu vào của hệ thống nhận dạng hoa là một bức ảnh được thu lại từ một máy ảnh hoặc các thiết bị ghi hình khác. Các ảnh này đôi khi có chất lượng không như mong muốn, ví dụ như quá tối, chủ thể quá nhỏ, quá to. Nguyên nhân chủ yếu có thể do thiết bị ghi hình không tốt hoặc điều kiện ghi hình không tốt ví dụ trời quá tối, chụp ảnh trong khi đang di chuyển, người chụp ảnh không chú ý. Do vậy để bước phân loại có được kết quả tốt nhất ta cần phải có một bước tiền xử lý ảnh.

Trong nghiên cứu [1] [2] các tác giả sử dụng phương pháp tách phông nền MRF [8]. Theo như hình ảnh dưới đây được lấy trong nghiên cứu [1] các ảnh bên trên là ảnh gốc và các ảnh phía dưới đã được tách phông nền. Ta có thể thấy ở bức ảnh thứ 2 bông hoa sau khi tách phông nền đã bị mất đi 1 phần, tỷ lệ mất mát đối tượng sau tách phông nền theo như báo cáo là 6%.



Hình 2.1: Hình ảnh trong nghiên cứu [1] về phương pháp tách phông nền

Bên cạnh đó còn một kỹ thuật nữa giúp cải thiện tập dữ liệu là tạo ra nhiều dữ liệu hơn từ dữ liệu ban đầu. Dối với xử lý ảnh, kỹ thuật này có tên là ảnh mở rộng (Image Augmentation) [11] [17]. Ví dụ như sau, chúng ta có một bông hoa hồng thì khi soi vào gương, dịch sang trái, phải, trên, dưới, đứng trong tối, ngoài sáng, phóng to, thu nhỏ, xoay 1 góc bất kỳ thì cũng vẫn là một bông hồng đó. Vậy thì từ một ảnh huấn luyện, ta có thể tạo ra vô vàn những ảnh tương tự và vẫn có thể được nhận ra là bông hồng bằng các phép biến đổi ảnh như: xoay ảnh, lật ảnh, dịch ảnh,.. Và như vậy, chúng ta có một tập dữ liệu lớn hơn cho bước huấn luyện. Nghiên cứu [4] có sử dụng phương pháp này và đã tăng được độ chính xác từ 74,4% lên thành 86.8%

Các nghiên cứu về nhận dạng dựa trên vector đặc trưng

Trích xuất thông tin là một phương pháp từ các dữ liệu thô ban đầu chuyển thành các tập, thuộc tính giúp chúng ta biểu diễn tập dữ liệu ban đầu tốt hơn. Cụ thể trong bài toán trích xuất thông tin từ ảnh ta có dữ liệu đầu vào là một hình ảnh, hay nói cách khác là một tập hợp biểu diễn các điểm ảnh. Tùy vào các mục đích khác nhau ta sẽ có những phương pháp trích xuất thuộc tính riêng biệt.

Nghiên cứu [1] sử dụng phương pháp trích xuất các đặc trưng phổ biến của một bức ảnh bao gồm màu sắc, khối hình và hoa văn từ đó sử dụng phương pháp láng giềng gần nhất để phân loại ra tên của loài hoa đó. Kết quả phân loại tốt nhất đạt được bằng cách kết hợp hình dạng và màu sắc đạt kết quả 81.3%

Bảng 2.1: Bảng kết quả của nghiên cứu [2]

Thuộc tính trích xuất	Độ chính xác phân loại (%)
HSV	43.0
SIFT internal	55.1
SIFT boundary	32.0
HOG	49.6
HSV + SIFT internal	66.4
HSV + SIFT boundary	57.0
HSV + HOG	62.1
SIFT internal + HOG	66.4
SIFT boundary + HOG	55.3
HSV + SIFT internal + HOG	71.8
HSV + SIFT internal + SIFT boundary + HOG	72.8

Bảng 2.2 cho thấy các kết quả của nghiên cứu [2], tại đó các tác giả đã tính toán bốn đặc điểm khác nhau cho các bông hoa, cụ thể là hình dạng kết cấu cục bộ, hình dạng của ranh giới, phân bố không gian tổng thể của cánh hoa và màu sắc của các loài hoa. Sau đó kết hợp các tính năng bằng cách sử dụng phân loại SVM. Kết quả cho thấy độ chính xác phân loại cho các bộ thuộc tính khác nhau. Có thể thấy rằng

việc kết hợp tất cả các thuộc tính mang lại độ chính xác tốt hơn nhiều so với sử dụng các thuộc tính đơn lẻ.

Vector đặc trưng về khối hình

Theo thứ tự từ trái qua phải: hoa gió, hoa buttercup và hoa thủy tiên. Nhìn qua thì các loài hoa này có cấu trúc tương tự nhau nhưng lại không hoàn toàn giống nhau. Sử dụng các mô tả SIFT [6] trên một lưới thông thường độ rộng của lưới M điểm ảnh, với phạm vi từ 10 đến 70 pixel, khu vực hỗ trợ cho Tính toán SIFT với bán kính R dao động từ 10 đến 70 pixel.



Hình 2.2: Hình ảnh trong nghiên cứu [1] về các loài hoa có hình dáng tương tự nhau

Vector đặc trưng về kết cấu trên vùng tiền cảnh sử dụng SIFT

Các mô tả SIFT được tính toán tại các điểm trên một lưới thông thường với khoảng cách M điểm ảnh trên vùng hoa tiền cảnh. Tại mỗi điểm lưới, các bộ mô tả được tính toán trên các cửa sổ hỗ trợ hình tròn với bán kính R điểm ảnh kết quả thu được là một vectơ 128 chiều.

Vector đặc trưng kết cấu trên vùng ranh giới sử dụng SIFT

Ranh giới của ảnh sau bước tiền xử lý tách phông nền cho ta ranh giới của hoa. Sau đó lấy mẫu các mô tả SIFT trên đường viền của bông hoa, chúng ta có thể nhấn mạnh hơn đến hình dạng cục bộ của đường biên. Sử dụng các cửa sổ hình tròn bán kính R điểm ảnh và bước nhảy S điểm ảnh đọc theo đường biên của bông hoa kết quả thu được cũng thu được một vector 128 chiều.

Vector đặc trưng về màu sắc

Một số hoa có màu sắc rất đặc trưng và có thể giúp thu hẹp đáng kể các loài

để phân loại, nhưng nó sẽ không cho phép chúng ta xác định chính xác các loài hoa. Sử dụng mô tả màu bằng không gian màu HSV để giảm hiệu ứng của ánh sáng môi trường tới ảnh.

Vector đặc trưng về hoa văn Một số hoa có hoa văn đặc trưng trên cánh hoa của chúng. Những họa tiết này có thể đặc biệt hơn, chẳng hạn các sọc của hoa păng xê, các ô vuông ở hoa đầu rắn, hoặc các chấm nhỏ ở hoa ly hổ



Hình 2.3: Hình ảnh trong nghiên cứu [1] về các loài hoa có hình dáng tương tự nhau

Sử dụng bộ lọc MR8 [12] với vùng hỗ trợ hình vuông có kích thước từ 3 đến 19 pixel. Kết quả trên cho thấy độ chính xác phân loại thay đổi theo số lượng cụm cho các kích thước bộ lọc khác nhau.

Vector đặc trưng về phân bố tổng thể

Sử dụng phương pháp HOG [7], tương tự như các tính năng SIFT, ngoại trừ việc chúng sử dụng tương phản cục bộ chồng chéo đã chuẩn hóa giữa các ô trong một lưới. Tuy nhiên, thay vì được áp dụng cho các vùng cục bộ (bán kính R trong trường hợp SIFT), HOG được áp dụng ở đây trên toàn bộ vùng hoa. Theo cách này, sẽ mô tả được sự phân bố không gian toàn cầu của hoa.

Nghiên cứu về trích xuất thông tin từ ảnh gốc

Mạng tích chập OverFeat [4] Trong nghiên cứu này [4] các tác giả sử dụng mạng tích chập có tên là OverFeat. Mỗi lớp tích chập có chứa 96 đến 1024 các cửa sổ tích chập có kích thước 3 x 3 đến 7 x 7. Chỉnh lưu nửa sóng được sử dụng làm chức năng

kích hoạt phi tuyến. Các hạt nhân gộp tối đa (Max polling) có kích thước 3 x 3 và 5 x 5 được sử dụng ở các lớp khác nhau. Mô hình nhận đầu vào là hình ảnh màu có kích thước 221 x 221. OverFeat đã được đào tạo cho nhiệm vụ phân loại hình ảnh của ImageNet ILSVRC 2013 và thu được kết quả khá cao.

Các tác giả sử dụng bộ dữ liệu ImageNet ILSVRC13 chứa 1,2 triệu hình ảnh được dán nhãn bằng tay của 1000 loại nhãn để huấn luyện cho mô hình. Sau đó có thử nghiệm trên bộ dữ liệu [2] và thu được kết quả như bảng sau.

Bảng 2.2: Bảng kết quả của nghiên cứu [4]

Phương pháp	Độ chính xác phân loại (%)
CNN-SVM w/o seg	74.7
CNNaug-SVM w/o seg	86.8

Đối với tất cả các thí nghiệm đều sử dụng đến lớp kết nối đầy đủ đầu tiên (lớp 22) của mạng để cho ra vector thuộc tính. Đối với tất cả các thử nghiệm, thay đổi kích thước toàn bộ hình ảnh thành 221 x 221. Đầu ra sẽ cho một vector thuộc tính có 4096 chiều. Trong nghiên cứu có đề cập tới hai loại cài đặt là có sử dụng phương pháp tăng cường ảnh CNNaug và không sử dụng phương pháp tăng cường ảnh. Đối với các nhu cầu phân loại trong đó các nhãn không loại trừ lẫn nhau các tác giả sử dụng chiến lược một với tất cả (one-against-all), và trong phần còn lại của các thí nghiệm sử dụng các SVM tuyến tính một đối một (one-against-one).

Mạng tích chập VGG19 [16]

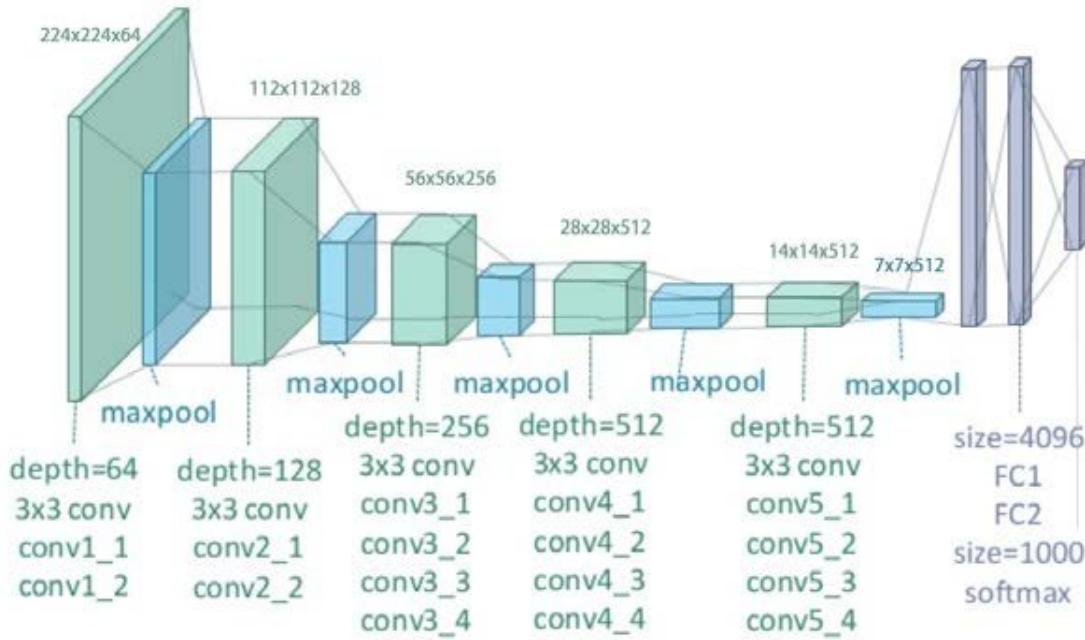
Dựa trên các công trình nghiên cứu trước đó thì vào năm 2015 có một nghiên cứu [16] đề xuất sử dụng mạng tích chập ConvNet [16] gồm rất nhiều lớp để có thể đáp ứng được khả năng phân loại trên diện rộng. Trong nghiên cứu này, các tác giả đã điều tra ảnh hưởng của độ sâu mạng tích chập đến độ chính xác trong việc nhận dạng hình ảnh với quy mô lớn.

Qua đánh giá kỹ lưỡng các mạng đã được tăng độ sâu bằng cách sử dụng kiến trúc với các bộ lọc tích chập rất nhỏ (3 x 3), cho thấy có thể đạt được sự cải thiện đáng kể về cấu hình trước khi nâng độ sâu lên 16 đến 19 lớp.

Dưới đây là mô tả cách cài đặt cho mạng ConvNet. Độ sâu của các cấu hình tăng từ trái (A) sang phải (E), khi nhiều lớp được thêm vào (các lớp được thêm vào được in đậm). Các tham số của lớp tích chập được ký hiệu là kích thước trường xác nhận kích thước của bộ dữ liệu của kênh. Kết quả thu được là một vector thuộc tính 4096 chiều.

ConvNet Configuration					
A	A-LRN	B	C	D	E
11 weight layers	11 weight layers	13 weight layers	16 weight layers	16 weight layers	19 weight layers
input (224×224 RGB image)					
conv3-64	conv3-64 LRN	conv3-64 conv3-64	conv3-64 conv3-64	conv3-64 conv3-64	conv3-64 conv3-64
maxpool					
conv3-128	conv3-128	conv3-128 conv3-128	conv3-128 conv3-128	conv3-128 conv3-128	conv3-128 conv3-128
maxpool					
conv3-256 conv3-256	conv3-256 conv3-256	conv3-256 conv3-256	conv3-256 conv3-256 conv1-256	conv3-256 conv3-256 conv3-256	conv3-256 conv3-256 conv3-256 conv3-256
maxpool					
conv3-512 conv3-512	conv3-512 conv3-512	conv3-512 conv3-512	conv3-512 conv3-512 conv1-512	conv3-512 conv3-512 conv3-512	conv3-512 conv3-512 conv3-512 conv3-512
maxpool					
conv3-512 conv3-512	conv3-512 conv3-512	conv3-512 conv3-512	conv3-512 conv3-512 conv1-512	conv3-512 conv3-512 conv3-512	conv3-512 conv3-512 conv3-512 conv3-512
maxpool					
FC-4096					
FC-4096					
FC-1000					
soft-max					

Hình 2.4: Bảng mô tả cài đặt cho mạng VGG19 trong nghiên cứu [16]



Hình 2.5: Hình ảnh mô tả kiến trúc của VGG19 lấy từ trang www.researchgate.net

Nghiên cứu về nhận dạng ảnh dựa trên tìm kiếm

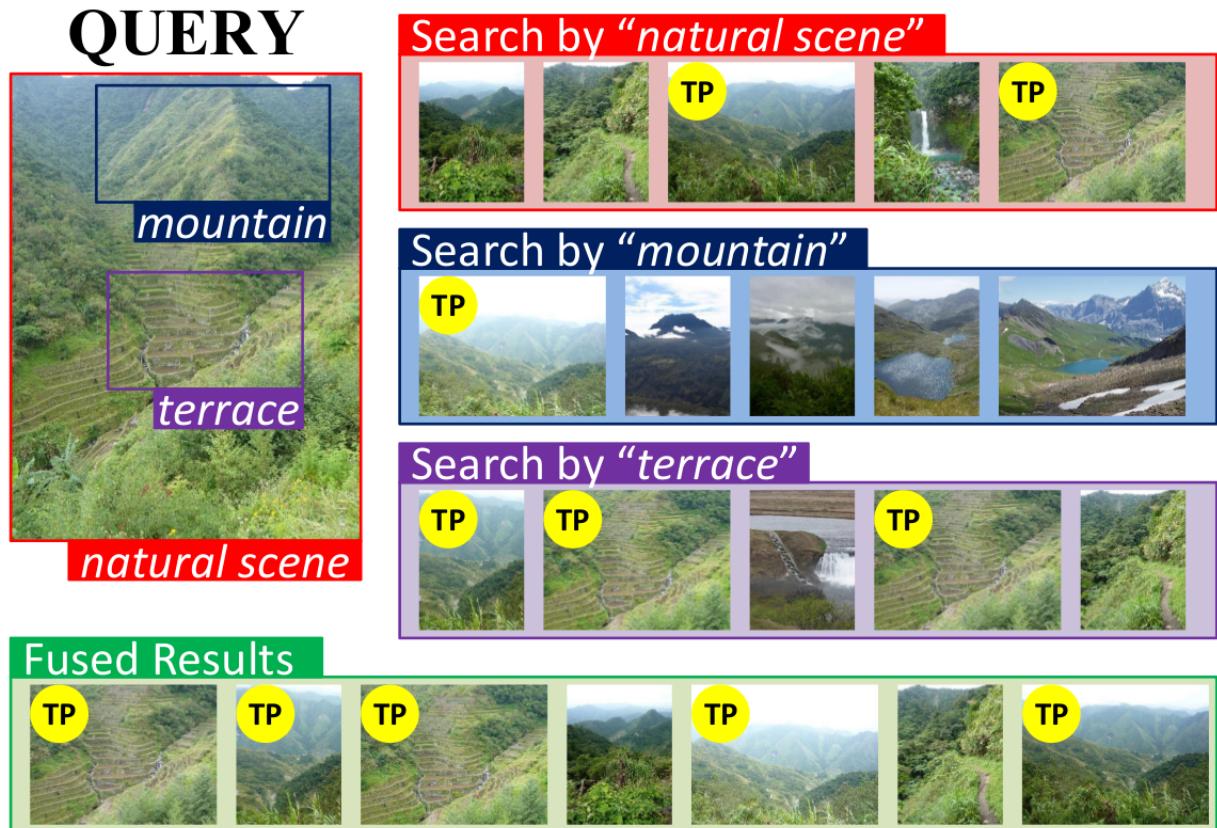
Phương pháp nhận dạng ảnh dựa trên tìm kiếm có chút khác biệt so với bài toán phân loại hình ảnh, đầu vào của bài toán vẫn là một hình ảnh và nhiệm vụ đặt ra ở đây là phải sắp xếp, hoặc tìm kiếm, những hình ảnh có liên quan trong cơ sở dữ liệu. Có hai hướng tiếp cận trong bài toán truy hồi là theo hướng khái niệm và hướng nội dung.

- Tiếp cận theo hướng khái niệm: là việc tìm kiếm dựa trên các thông tin liên quan đến một bức ảnh như tên bức ảnh, tiêu đề, chủ đề hoặc phần chữ miêu tả về bức ảnh đó.
- Tiếp cận theo hướng nội dung: là việc tìm kiếm dựa trên nội dung của ảnh.

Nghiên cứu về nhận dạng hoa có sử dụng phương pháp tìm kiếm hình ảnh [3] tại đây các tác giả sử dụng thuật toán ước lượng láng giềng gần nhất trực tiếp (Online Nearest neighbor Estimation - ONE) cho cả phân loại và truy hồi hình ảnh, bằng cách tính độ tương tự giữa các đối tượng trong ảnh dùng để tìm kiếm với từng lớp hoặc

các hình ảnh của lớp đó. Kết quả phân loại tốt nhất mà các tác giả đạt được trên bộ dữ liệu [2] là 86.24%

Các bước thực hiện của phương pháp



Hình 2.6: Hình ảnh tóm tắt về phương pháp ước lượng láng giềng gần nhất trong nghiên cứu [3] của tác giả

Đầu tiên sử dụng phương pháp phát hiện đa vật thể dựa theo các nghiên cứu [13] [14] [15] trên hình ảnh truy vấn và từng hình ảnh của các lớp. Ở đây hình ảnh bên trái ở trên là hình ảnh cần tìm kiếm ta sẽ sử dụng phương pháp phát hiện đa vật thể ra được các vật thể có trong bức ảnh là "mountain", "terrace" và "natural scene".

Trong bộ dữ liệu của ta, ta sẽ lấy ra các lớp về "mountain", "terrace" và "natural scene". Rồi sau đó trích xuất các thuộc tính trên mỗi đối tượng đó để cung cấp một mô tả sát nhất cho hình ảnh tìm kiếm.

So sánh mức độ liên quan giữa các đối tượng đã được phát hiện trong ảnh truy

vẫn với các đối tượng trong các ảnh thành viên trong lớp đó. Các ảnh nào có độ tương tự gần nhất với các đối tượng có mặt trong hình ảnh cần tìm kiếm sẽ có chữ "TP - true positive" màu vàng.

Trong bước tìm kiếm, mức độ liên quan của hình ảnh cần tìm kiếm đối với một danh mục hoặc hình ảnh ứng cử viên được tính bằng mức trung bình gần nhất khoảng cách từ hình ảnh cần tìm kiếm đối tượng đến các đối tượng trong lớp đó.

Các ứng dụng di động cho nhận dạng hoa

Các ứng dụng nhận dạng hoa và thực vật có xuất hiện khá nhiều nhưng lại có sự khác biệt khá lớn về độ chính xác và tốc độ nhận dạng. Trong khóa luận này lấy ra 3 ví dụ về ứng dụng nhận dạng hoa có độ chính xác và tốc độ nhận dạng khá tốt.

- PictureThis - Flower & Plant Identification: do công ty Glority Software Limited.
- PlantNet Plant Identification: do công ty plantnet-project.org.
- PlantFinder - Flower & Plant Identification: do công ty HD Wallpapers Production.



Hình 2.7: Hình ảnh đại diện của các ứng dụng

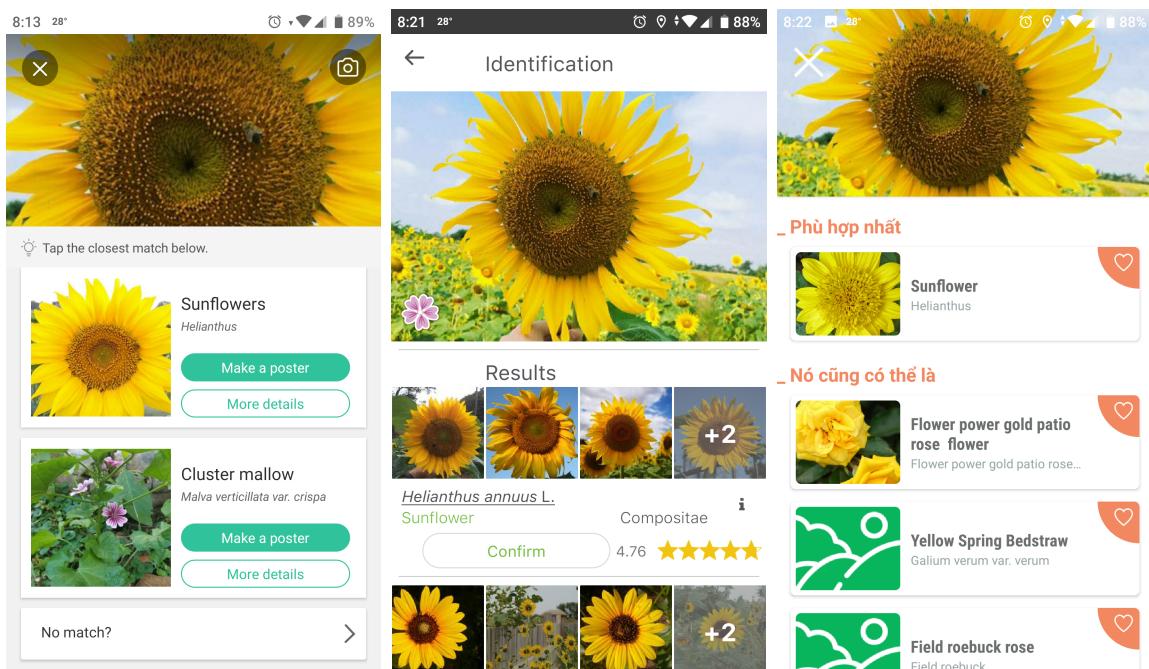
(Lần lượt từ trái qua phải ta có các ứng dụng là PictureThis, PlantNet, PlantFinder)
Ưu điểm chung của các ứng dụng:

- Các ứng dụng có độ nhận dạng chính xác khá tốt.

- Thời gian phản hồi ở mức khá.
- Có nêu ra đặc điểm chi tiết của loài hoa.

Một số nhược điểm còn tồn tại:

- Toàn bộ tên loài hoa và chi tiết về đặc điểm của loài hoa đang ở tiếng anh và tiếng latin, gây ra đôi chút khó khăn cho người Việt khi cần tra cứu.
- Một số ứng dụng sau khi chụp xong cần phải khoanh vùng hoa cần tìm bằng tay, chưa tự động phát hiện.
- Một số ứng dụng chưa có bước phát hiện xem ảnh có hoa hay không.

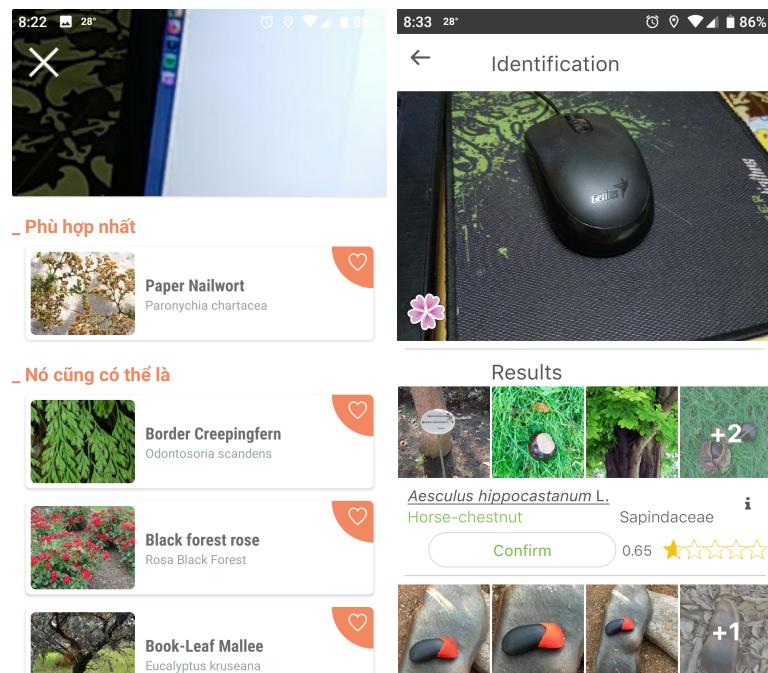


Hình 2.8: Hình ảnh kết quả toàn tiếng Anh khi tra cứu ảnh hoa hướng dương bằng các ứng dụng

(Lần lượt từ trái qua phải là ảnh chụp màn hình kết quả của các ứng dụng PlantFinder, PlantNet và cuối cùng là Picture this.)



Hình 2.9: Hình ảnh về ứng dụng PlantFinder cần phải khoanh vùng phần có hoa bằng tay



Hình 2.10: Hình ảnh về ứng dụng chưa có chức năng phát hiện hoa trong ảnh

Chương 3

Phát triển phần mềm nhận dạng hoa trên thiết bị di động

Như đã trình bày tại chương 2, bước trích xuất thông tin từ ảnh là bước rất quan trọng trong mô hình nhận diện ảnh làm tốt bước này ta sẽ cải thiện được đáng kể độ chính xác và hiện nay có rất nhiều phương pháp để trích xuất thông tin từ ảnh. Để tăng cường độ chính xác nhận diện ảnh cho ứng dụng tôi có thêm một bước vào trước bước phân loại hoa đó là phát hiện ảnh có hoa hay không.

Để có thể làm tốt bước phát hiện hoa ta cần phải trích xuất được thuộc tính của cả các hình ảnh mà không có hoa trong đó, các hình ảnh không có hoa trong đó có thể thuộc rất nhiều các chủ đề như phong cảnh, động vật, đường phố Trong nghiên cứu [4] được công bố năm 2014 các tác giả đã đề xuất sử dụng mạng tích chập đã được huấn luyện bởi 1.2 triệu hình ảnh được lấy từ bộ dữ liệu ImageNet các hình ảnh thuộc rất nhiều chủ đề. Và kết quả thu được khi sử dụng một mô hình trích xuất nhưng lại có thể đáp ứng cho nhiều chủ đề ảnh khác nhau và thu được kết quả khá khả quan đối với chính bộ dữ liệu Oxford-102.

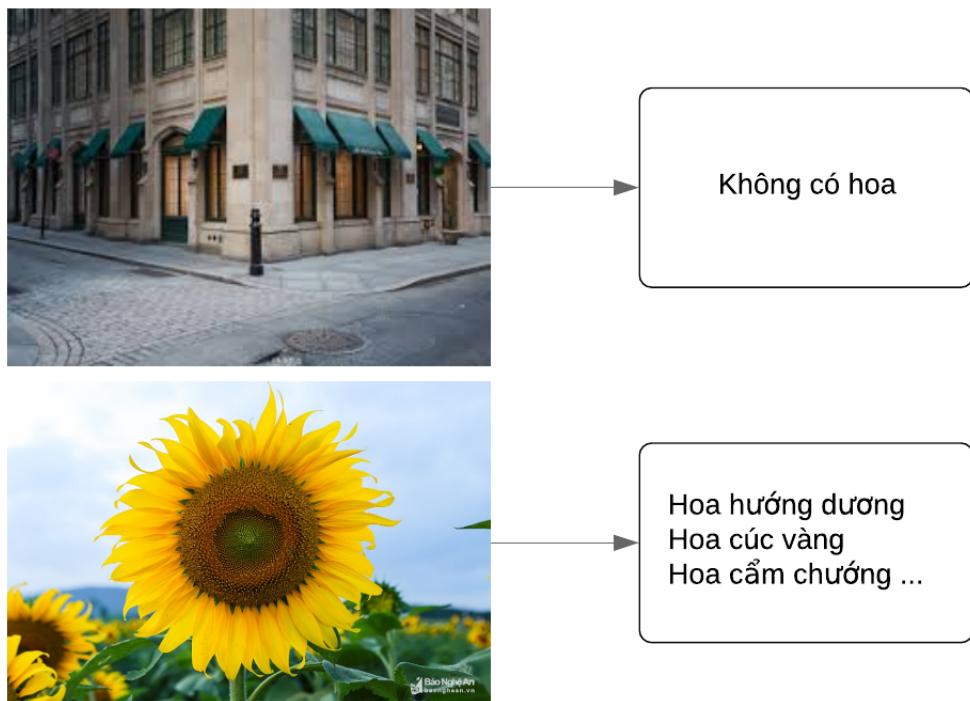
Theo nghiên cứu [16] cũng sử dụng mạng tích chập nhưng với mô hình có độ sâu

các tầng tích chập lớn hơn thì sẽ cho được kết quả khả quan hơn. Vì vậy tôi đề xuất sử dụng mô hình mạng tích chập VGG19 [16] đã được xây dựng sẵn để làm bước trích xuất thông tin cho tất cả các bước trích xuất thông tin trong mô hình hệ thống.

Mô tả bài toán

Ứng dụng nhận dạng hoa trong khóa luận này sẽ có chức năng chính đó là nhận dạng hoa thông qua ảnh chụp. Như vậy đầu vào và đầu ra của bài toán sẽ là một hình ảnh được ghi lại từ các thiết bị ghi hình và đầu ra có thể có hai trường hợp

- Trường hợp 1: Trong ảnh không có hoa thì kết quả trả về là không tồn tại hoa.
- Trường hợp 2: Trong ảnh có hoa thì kết quả trả về là một danh sách 5 loài hoa có kết quả nhận dạng tốt nhất.

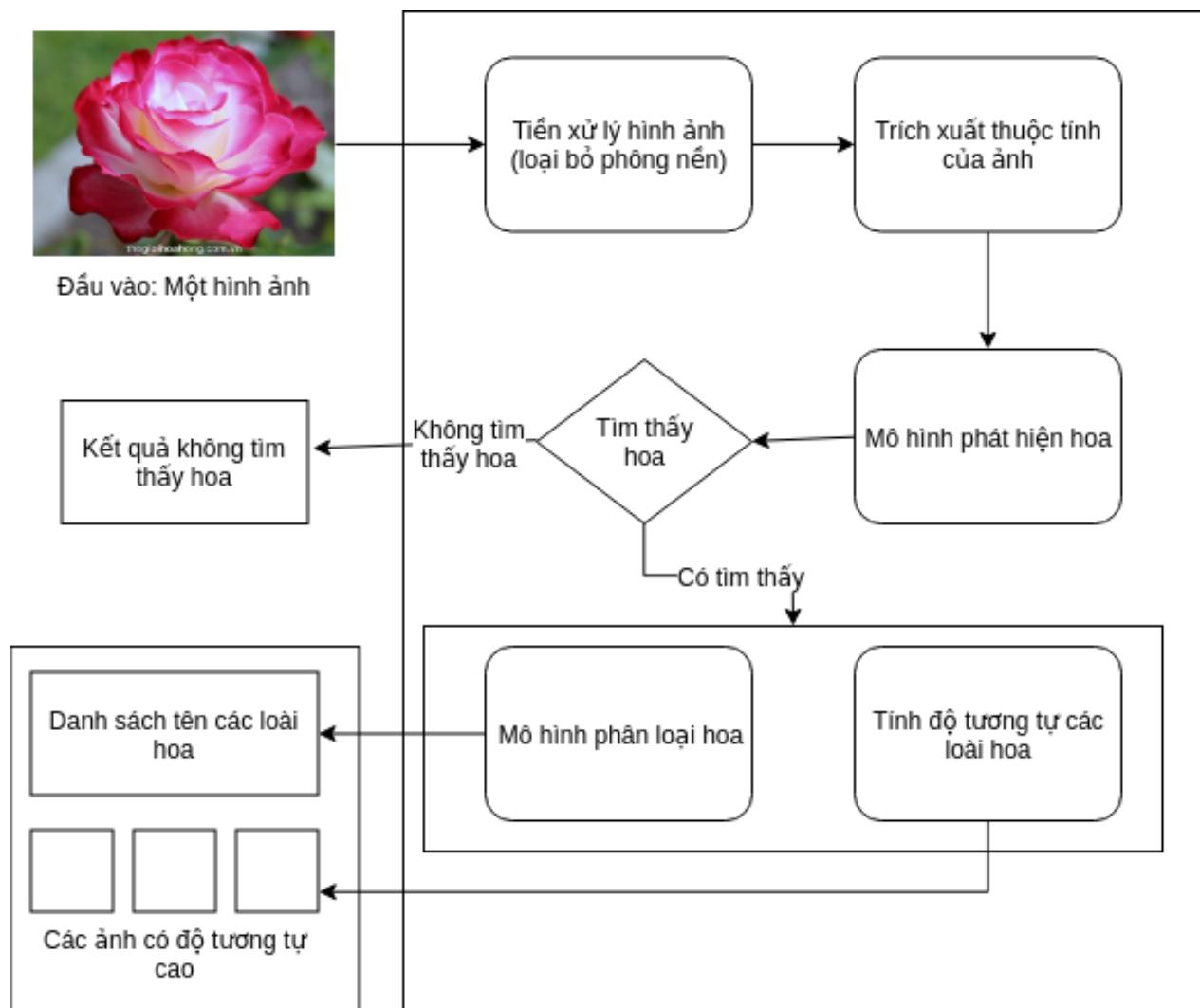


Hình 3.1: Ví dụ về đầu vào và đầu ra của bài toán

Mô tả mô hình hệ thống

Dưới đây là mô tả và giải thích chi tiết về các bước trong mô hình hệ thống qua sự tham khảo các nghiên cứu [1] [2] [3] [4] [16] và học hỏi về điểm mạnh và điểm yếu của các ứng dụng đã có về nhận dạng hoa trên thiết bị di động.

Mô hình chung



Hình 3.2: mô hình chung của hệ thống

Bước 1: Tiền xử lý dữ liệu.

- Do các ảnh của người dùng tải lên hệ thống có thể gấp phải một số trường hợp đặc biệt ví dụ như ảnh chụp bông hoa quá nhỏ, ảnh của hoa không nằm trong chính giữa bức ảnh ... Vì vậy, ảnh sau khi được chụp và gửi lên hệ thống sẽ được tiền xử lý bằng phương pháp tách phông nền để loại bỏ các phần không quan trọng.
- **Đầu vào:** Ảnh do người dùng chụp gửi lên hệ thống
- **Đầu ra:** Ảnh đã được tách phông nền.



Hình 3.3: ví dụ về bước tách phông nền

Bước 2: Trích xuất thuộc tính.

- Đầu vào của bước này đó là bức ảnh mới có trọng tâm là chủ thể cần nhận dạng.

- Ảnh sẽ được thay đổi kích cỡ về dạng 224 x 224.
- Sử dụng mô hình VGG19 [16] để làm bước trích xuất thuộc tính. Dựa trên các cài đặt của mô hình VGG19 [16] ta đã nói ở phần 2.2.3 trong chương 2. Chỉ sử dụng đến tầng thứ 19 FC-4096 của mô hình và kết quả thu được ở đây là một vector thuộc tính có 4096 chiều.
- Vector thuộc tính sẽ được chuẩn hóa sẽ được chuẩn hóa L2 (L2 normalization).

```

1   from keras.applications.vgg19 import VGG19
2   from keras.preprocessing import image
3   from keras.applications.vgg19 import preprocess_input
4
5   base_model = VGG19(weights='imagenet')
6   model = Model(inputs=base_model.input,
7                 outputs=base_model.get_layer('fc2').output)
8
9   def get_feature_1_image(image_name):
10      img_path = './' + image_name + '.jpg'
11      img = image.load_img(img_path, target_size=(224, 224))
12      x = image.img_to_array(img)
13      x = np.expand_dims(x, axis=0)
14      x = preprocess_input(x)
15      features = model.predict(x)

```

Bước 3: Phát hiện hoa trong ảnh.

- Sử dụng phương pháp máy vector hỗ trợ để xây dựng mô hình phân loại nhị phân (binary classification).
- Đầu vào của bước này là vector thuộc tính đã được sinh ra ở bước 2.
- Nếu kết quả của mô hình phát hiện ảnh là không có hoa thì lập tức trả lại kết quả cho người dùng là không tìm thấy hoa.
- Nếu kết quả là có hoa thì chuyển vector thuộc tính sang bước 4.

Bước 4: Phân loại hoa.

- Sử dụng phương pháp máy vector hỗ trợ để xây dựng mô hình phân loại chi tiết về cách xây dựng mô hình sẽ được nói kỹ hơn ở phần sau của chương.
- Đầu vào của bước này là vector thuộc tính đã qua bước phát hiện hoa.
- Kết quả của mô hình phân loại ở bước này sẽ cho ta danh sách tên các loài hoa theo thứ tự là có độ phù hợp cao nhất đến thấp nhất. Trong ứng dụng này sẽ lấy ra tên 5 loài hoa có độ phù hợp cao nhất.

Như đã có ở bước thứ 4, kết quả đi ra sẽ là danh sách tên 5 loài hoa. Ở đây, nếu nhiệm vụ của bài toán chỉ là cho ra tên của loài hoa thì có vẻ như không cần bước này. Trong thực tế một loài hoa hồng như có thể xuất hiện trong tự nhiên với khá nhiều màu sắc như hoa hồng đỏ, hoa hồng trắng, vàng. Vì vậy khi tìm kiếm một bông hồng trắng nếu ta sẽ đưa ra danh sách tên của loài hoa đó và kèm theo đó là một bức ảnh có độ tương tự gần giống nhất với bức ảnh cần tìm kiếm của mỗi loài.

Bước 5: Tính độ tương tự của các loài.

- Độ tương tự được tính ở đây là cosin giữa hai vector. Ở trên đây là công thức tính độ tương tự giữa góc giữa hai hình ảnh A và B trong đó A và B lần lượt là các vector thuộc tính của 2 hình ảnh A và B.

$\|A\|$ là độ dài của vector thuộc tính A

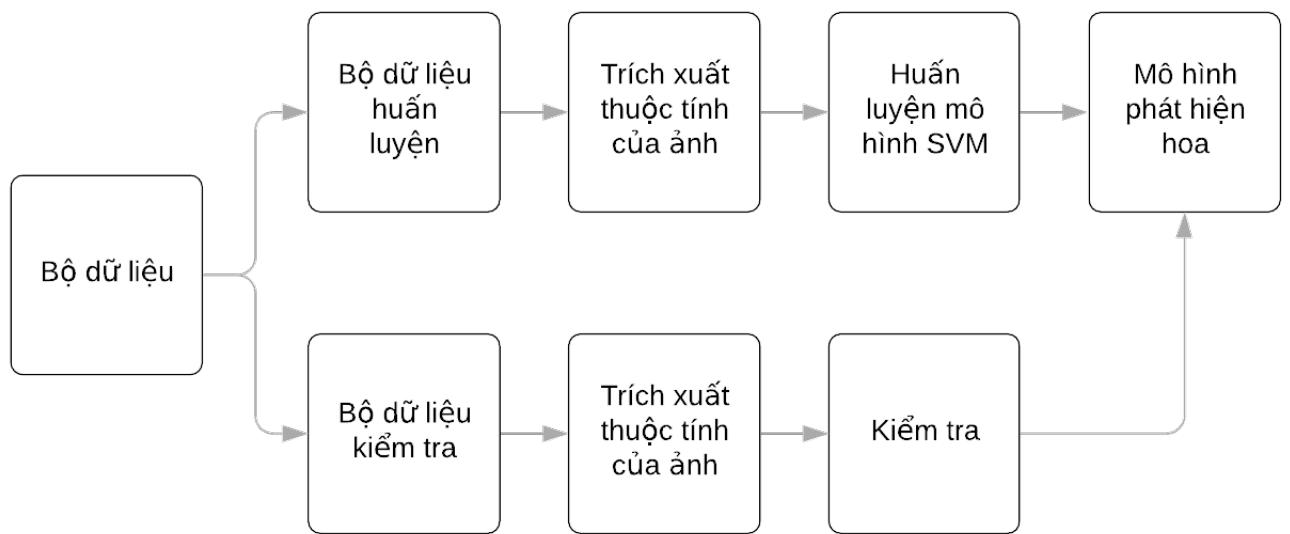
$\|B\|$ là độ dài vector thuộc tính B

Độ tương tự sẽ lớn nhất khi vector A trùng vector B và công thức trên cho kết quả bằng 1.

$$\text{similarity} = \cos(\theta) = \frac{\mathbf{A} \cdot \mathbf{B}}{\|\mathbf{A}\| \|\mathbf{B}\|} = \frac{\sum_{i=1}^n A_i B_i}{\sqrt{\sum_{i=1}^n A_i^2} \sqrt{\sum_{i=1}^n B_i^2}},$$

- Trong mỗi loài hoa được nhận ra ảnh nào có độ tương tự lớn nhất với ảnh cần tìm thì ta sẽ hiển thị ra ảnh đó, để người dùng có thể dễ dàng xác nhận được đúng loài hoa cần tìm.

Mô hình phát hiện hoa



Hình 3.4: Mô hình phát hiện hoa

Bước 1: Chuẩn bị dữ liệu.

- Bộ dữ liệu bao gồm 10.290 ảnh trong đó có 5132 ảnh có hoa và 5158 ảnh được thu thập từ bộ dữ liệu ImageNet.
- Chia bộ dữ liệu thành các tập con.
 - Bộ dữ liệu kiểm tra : 1000 ảnh (mỗi lớp có 500 ảnh).
 - Bộ dữ liệu xác nhận : 1000 ảnh (mỗi lớp có 500 ảnh).
 - Bộ dữ liệu huấn luyện: các ảnh còn lại (cụ thể 4132 ảnh có hoa và 4158 ảnh không có hoa).

Bước 2: Trích xuất thuộc tính.

- Ảnh sẽ được thay đổi kích cỡ về dạng 224 x 224.
- Sử dụng mô hình VGG19 [16] để làm bước trích xuất thuộc tính. Dựa trên các cài đặt của mô hình VGG19 [16] ta đã nói ở phần 2.2.3 trong chương 2. Chỉ sử dụng đến tầng thứ 19 FC-4096 của mô hình và kết quả thu được ở đây là một vector thuộc tính có 4096 chiều.
- Vector thuộc tính sẽ được chuẩn hóa sẽ được chuẩn hóa L2 (L2 normalization).

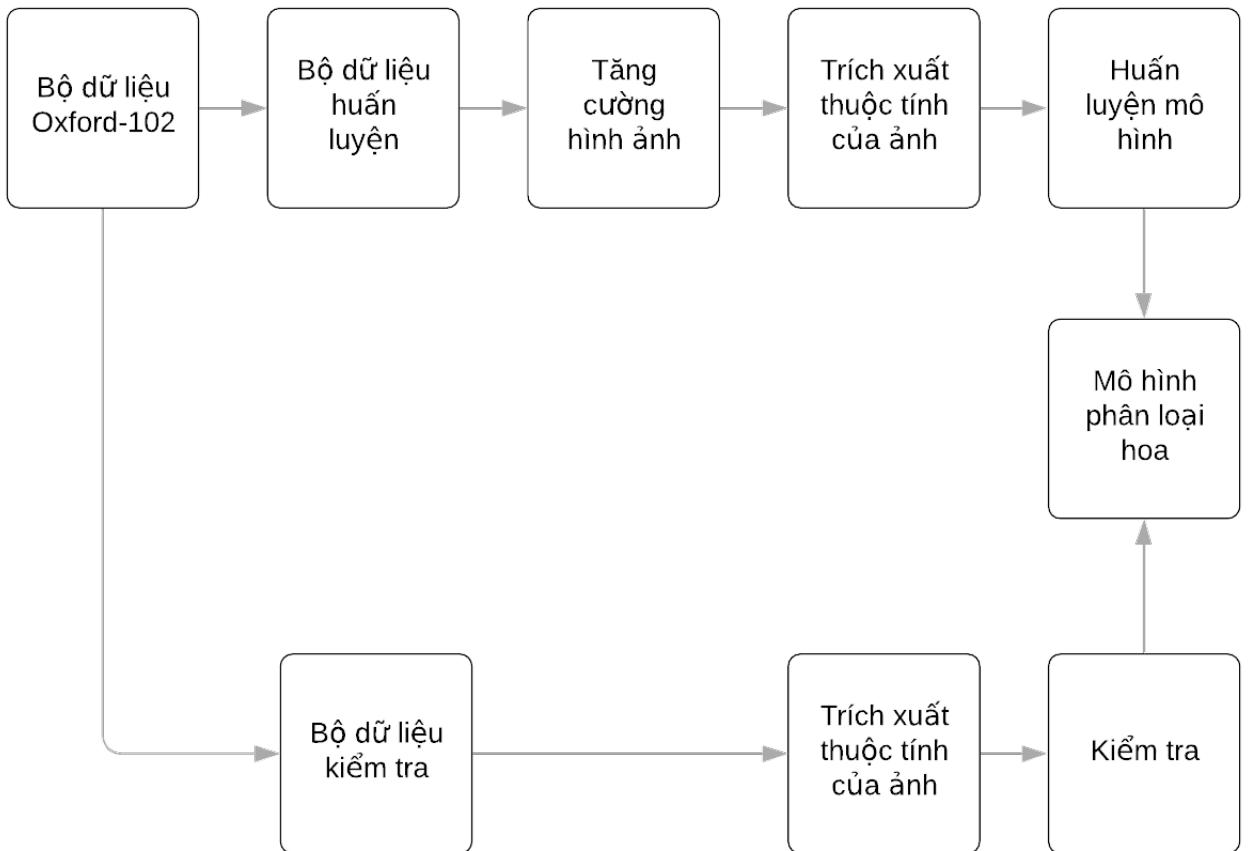
Bước 3: Huấn luyện mô hình.

- Sử dụng phương pháp máy vector hỗ trợ để xây dựng mô hình phân loại nhị phân (binary classification).
- Lấy bộ dữ liệu huấn luyện để huấn luyện cho mô hình, và sử dụng bộ xác nhận để đo độ chính xác của mô hình.
- Kết quả thu được là mô hình phát hiện hoa.

Bước 4: Kiểm tra mô hình.

- Sử dụng bộ kiểm tra để đo độ chính xác của mô hình vừa nhận được.

Mô hình phân loại hoa



Hình 3.5: Mô hình phân loại hoa

Bước 1: Chuẩn bị dữ liệu.

- Ở đây sử dụng bộ dữ liệu Oxford-102 [1], nhận thấy số lượng ảnh để huấn luyện cho mô hình chỉ là 10 ảnh mỗi loài hoa, như vậy là hơi ít ảnh vì vậy tôi đã gộp bộ huấn luyện và bộ xác nhận lại thành bộ huấn luyện mở rộng. Sau đó sử dụng phương pháp tăng cường ảnh [17] để tăng số lượng ảnh huấn luyện.
- Do bộ dữ liệu chia lượng ảnh huấn luyện ít mà ảnh dùng để kiểm tra lại nhiều nên tôi đã chia bộ dữ liệu theo cách mới đó là cắt một nửa bộ kiểm tra (mỗi

loài hoa trong tập kiểm tra chia một nửa cho bộ huấn luyện nửa còn lại để kiểm tra).

- Các ảnh trong bộ dữ liệu đều là ảnh chụp cận cảnh bông hoa vì vậy ở mô hình này không cần đến bước tách phông nền.

Bước 2: Trích xuất thuộc tính.

- Đầu vào của bước này đó là bức ảnh mới có trọng tâm là chủ thể cần nhận dạng.
- Ảnh sẽ được thay đổi kích cỡ về dạng 224 x 224.
- Sử dụng mô hình VGG19 [16] để làm bước trích xuất thuộc tính. Dựa trên các cài đặt của mô hình VGG19 [16] ta đã nói ở phần 2.2.3 trong chương 2. Chỉ sử dụng đến tầng thứ 19 FC-4096 của mô hình và kết quả thu được ở đây là một vector thuộc tính có 4096 chiều.
- Vector thuộc tính sẽ được chuẩn hóa sẽ được chuẩn hóa L2 (L2 normalization).

Bước 3: Huấn luyện mô hình.

- Sử dụng phương pháp máy vector hỗ trợ để xây dựng mô hình phân loại.
- Lấy bộ dữ liệu huấn luyện để huấn luyện cho mô hình. Kết quả thu được là mô hình phát hiện hoa.

Bước 4: Kiểm tra mô hình.

- Sử dụng bộ kiểm tra để đo độ chính xác của mô hình vừa nhận được.

Chương 4

Kết quả thực nghiệm

Tại chương 4 này sẽ nói về cách thu thập các bộ dữ liệu, triển khai cài đặt các phần của ứng dụng dựa trên các phương pháp giải quyết vấn đề đã đặt ra ở chương 3. Chúng ta có một ứng dụng nhận dạng các loài hoa sử dụng thư viện React Native viết cho cả hai hệ điều hành là android và IOS. Ứng dụng sẽ truyền và nhận dữ liệu từ một máy chủ sử dụng các mô hình phân loại đã được huấn luyện. Đồng thời trong chương này cũng công bố các kết quả về các mô hình phân loại đã đạt được

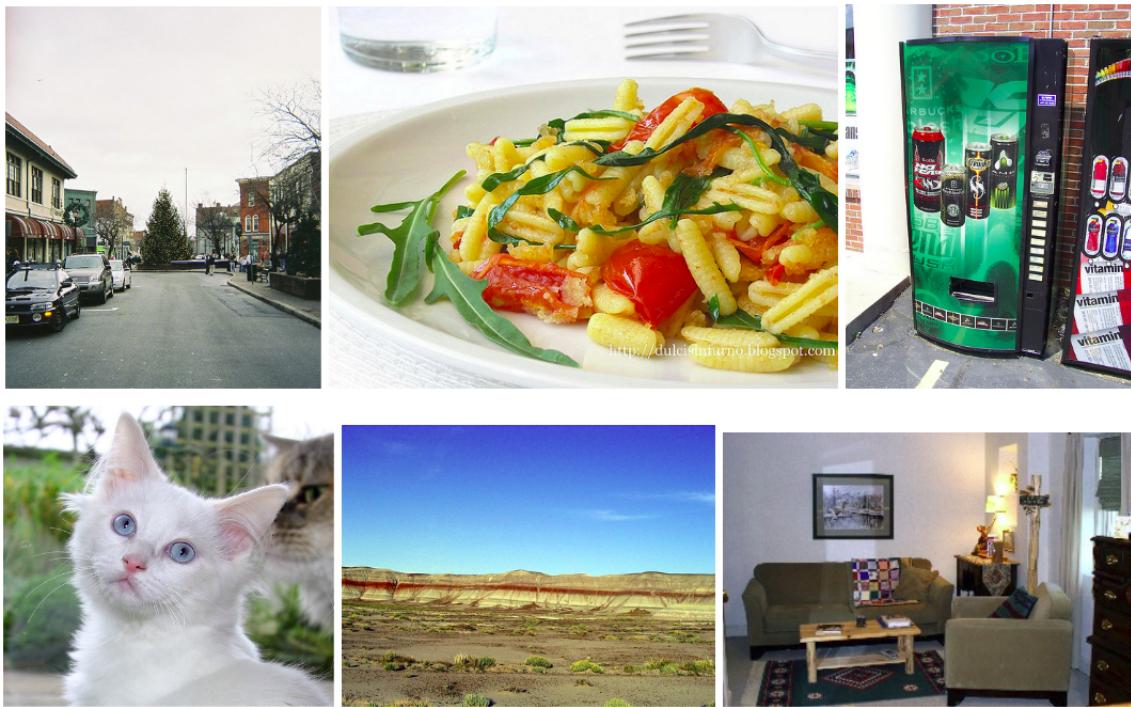
Xây dựng bộ dữ liệu.

Bộ dữ liệu dùng để phát hiện đối tượng hoa trong ảnh.

Cách thu thập: Sử dụng bộ dữ liệu ImageNet tại trang web : <http://www.image-net.org> để thu thập các ảnh. Các ảnh tại đây đã được chia sẵn thành các chủ đề nhỏ, đa dạng về chủ đề. Tổng cộng là 5158 ảnh không có hoa. Để lấy các bức ảnh không có hoa là chủ thể chính trong ảnh tôi lựa chọn các chủ đề:

- Động vật có vú : bao gồm 743 ảnh.
- Cảnh vật trong tự nhiên : bao gồm 930 ảnh.
- Đồ ăn : bao gồm 1063 ảnh.

- Cảnh vật trong phòng : bao gồm 659 ảnh.
- Công cụ máy móc : bao gồm 799 ảnh.
- Đường phố : bao gồm 964 ảnh.



Hình 4.1: Hình ảnh lấy trong bộ ảnh không có hoa

Bộ dữ liệu có hoa tôi lấy từ chủ đề "flower" thu thập được 1207 ảnh. Nhận thấy có sự chênh lệch khá lớn giữa bộ ảnh có hoa và không có hoa nên tôi sử dụng một phần bộ dữ liệu Oxford-102 [1], vào để số lượng ảnh có hoa và không có hoa mang tính cân đối. Tổng cộng là 5132 ảnh trong đó có:

- Ảnh có hoa từ ImageNet: bao gồm 1207 ảnh.
- Ảnh lấy từ Oxford-102 : bao gồm 3925 ảnh.

Chương 5

Kết luận

Trong khóa luận này trình bày một ứng dụng nhận dạng tên loài hoa dựa vào ảnh đầu vào trên nền tảng di động. Ứng dụng này được xây dựng với ba mô hình học máy chính: (1) một mô hình phân loại nhị phân để phát hiện có đối tượng hoa trong ảnh hay không, (2) một mô hình nhận dạng tên của loài hoa trong ảnh, và (3) một thuật toán tìm kiếm ảnh tương tự khi hiển thị các ảnh kết quả.

Ứng dụng được phát triển sử dụng phương pháp sử dụng lại một mô hình học sâu đã được huấn luyện trên bộ dữ liệu lớn khác để trích xuất dữ liệu của ảnh VGG19 [16]. Thực nghiệm so sánh phương pháp tiếp cận này với các cách trích xuất thuộc tính truyền thống chỉ ra tính khả thi và tiết kiệm của việc sử dụng lại mô hình giữa các bài toán nghiên cứu trong lĩnh vực xử lý ảnh.

Những đóng góp chính của khóa luận là: (1) Việt hoá tên cũng như đặc điểm về sinh trưởng, cách trồng của 102 loài hoa trong bộ dữ liệu Oxford-102, (2) phát triển ứng dụng trên nền tảng di động, và (3) phát triển mô hình phân loại ảnh có hoa hay không với độ chính xác 98.6% và mô hình nhận dạng tên loài hoa với độ chính xác là 78.56%.

Hướng phát triển tiếp theo của nghiên cứu là tìm cách khắc phục các điểm còn hạn chế trong việc nhận dạng hoa như nhận dạng đa chủ thể hoa hay nhận dạng hoa khi mọc cả chùm.

Tài liệu tham khảo

Tiếng Anh

- [1] Maria-Elena Nilsback and Andrew Zisserman. A Visual Vocabulary for Flower Classification Robotics Research Group, Department of Engineering Science University of Oxford, United Kingdom
- [2] Maria-Elena Nilsback and Andrew Zisserman. Automated flower classification over a large number of classes Visual Geometry Group, Department of Engineering Science University of Oxford, United Kingdom
- [3] Lingxi Xie , Richang Hong , Bo Zhang , and Qi Tian. Image Classification and Retrieval are ONE
- [4] CNN Features off-the-shelf: an Astounding Baseline for Recognition Ali Sharif Razavian Hossein Azizpour Josephine Sullivan Stefan Carlsson CVAP, KTH (Royal Institute of Technology) Stockholm, Sweden
- [5] M. Varma and D. Ray. Learning the discriminative power invariance trade-off. In Proc. ICCV, 2007.
- [6] D. Lowe. Distinctive image features from scale-invariant keypoints. IJCV, 60(2):91–110, 2004.
- [7] N. Dalal and B. Triggs. Histogram of oriented gradients for human detection. In Proc. CVPR, volume 2, pages 886–893, 2005.

- [8] Y. Y. Boykov and M. P. Jolly. Interactive graph cuts for optimal boundary and region segmentation of objects in N-D images. In Proc. ICCV, volume 2, pages 105– 112, 2001.
- [9] Luca Bertelli. Tianli Yu. Diem Vu Burak Gokturk. Kernelized Structural SVM Learning for Supervised Object Segmentation. Google, Inc.
- [10] Luca Bertelli. Tianli Yu. Diem Vu Burak Gokturk. Kernelized Structural SVM Learning for Supervised Object Segmentation. Google, Inc.
- [11] L. Perez and J. Wang. The effectiveness of data augmentation in image classification using deep learning. arXiv preprint arXiv:1712.04621, 2017
- [12] M. Varma and A. Zisserman. Classifying images of materials: Achieving viewpoint and illumination independence. In Proc. ECCV, volume 3, pages 255–271. Springer-Verlag, May 2002.
- [13] B. Alexe, T. Deselaers, and V. Ferrari. Measuring the Objectness of Image Windows. TPAMI, 2012.
- [14] J. Uijlings, K. van de Sande, T. Gevers, and A. Smeulders. Selective Search for Object Recognition. IJCV, 2013.
- [15] M. Cheng, Z. Zhang, W. Lin, and P. Torr. BING: Binarized Normed Gradients for Objectness Estimation at 300fps. CVPR, 2014.
- [16] K. Simonyan and A. Zisserman. Very deep convolutional networks for large-scale image recognition. In ICLR, 2015.
- [17] Marcus D. Bloice, Christof Stocker, Andreas Holzinger. Augmentor: An Image Augmentation Library for Machine Learning. Medical University of Graz Graz, Austria